

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/242,657B

DATE: 02/12/2001

TIME: 10:52:56

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02122001\I242657B.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

4 (i) APPLICANT: Peter Ruhdal Jensen

5 Karin Hammer

7 (ii) TITLE OF INVENTION: Artificial promoter libraries

8 for selected organisms and promoters derived from such  
9 libraries

11 (iii) NUMBER OF SEQUENCES: 58

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: Stanislaus Aksman  
15 Hunton & Williams

16 (B) STREET: 1900 K Street, NW

17 (C) CITY: Washington, DC

18 (E) COUNTRY: USA

19 (F) ZIP: 20006-1109

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk

23 (B) COMPUTER: IBM PC compatible

24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30  
26 (EPO)

28 (vi) CURRENT APPLICATION DATA:

29 (A) APPLICATION NUMBER: US/09/242,657B

C--&gt; 30 (B) FILING DATE: 19-Feb-1999

32 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: DK 886/96

34 (B) FILING DATE: 23-AUG-1996

35 (A) APPLICATION NUMBER: PCT/DK97/00342

36 (B) FILING DATE: August 25, 1997

38 (viii) ATTORNEY/AGENT INFORMATION:

39 (C) REFERENCE/DOCKET NUMBER: 55411.000002

41 (ix) TELECOMMUNICATION INFORMATION:

42 (A) TELEPHONE: (202) 955-1500

43 (B) TELEFAX: (202) 778-2201

45 (2) INFORMATION FOR SEQ ID NO: 1:

47 (i) SEQUENCE CHARACTERISTICS:

48 (A) LENGTH: 100 base pairs

49 (B) TYPE: nucleic acid

50 (C) STRANDEDNESS: double

51 (D) TOPOLOGY: linear

53 (ii) MOLECULE TYPE: DNA (genomic)

55 (iii) HYPOTHETICAL: YES

56 (iv) ANTI-SENSE: NO

58 (vi) ORIGINAL SOURCE:

59 (A) ORGANISM: Lactococcus lactis

61 (ix) FEATURE:

62 (A) NAME/KEY: promoter

RAW SEQUENCE LISTING DATE: 02/12/2001  
PATENT APPLICATION: US/09/242,657B TIME: 10:52:56

Input Set : A:\Pto.amc  
Output Set: N:\CRF3\02122001\I242657B.raw

63 (B) LOCATION:26..82  
64 (C) IDENTIFICATION METHOD: experimental  
65 (D) OTHER INFORMATION:/evidence= EXPERIMENTAL  
66 /standard\_name= "Artificial promoter library" /note= "A  
67 degenerated sequence specifying a mixture of artificial  
68 promoters covering a wide range of expression in small steps  
69 in L. lactis"  
71 (ix) FEATURE:  
72 (A) NAME/KEY: misc\_feature  
73 (B) LOCATION:31..45  
74 (D) OTHER INFORMATION:/standard\_name= "Consensus  
75 sequence"  
77 (ix) FEATURE:  
78 (A) NAME/KEY: misc\_feature  
79 (B) LOCATION:60..69  
80 (D) OTHER INFORMATION:/standard\_name= "Consensus  
81 sequence"  
83 (ix) FEATURE:  
84 (A) NAME/KEY: misc\_feature  
85 (B) LOCATION:74..82  
86 (D) OTHER INFORMATION:/standard\_name= "Consensus  
87 sequence"  
89 (ix) FEATURE:  
90 (A) NAME/KEY: -35\_signal  
91 (B) LOCATION:40..45  
92 (D) OTHER INFORMATION:/standard\_name= "-35 box"  
94 (ix) FEATURE:  
95 (A) NAME/KEY: -10\_signal  
96 (B) LOCATION:63..68  
97 (D) OTHER INFORMATION:/standard\_name= "Pribnow box"  
99 (ix) FEATURE:  
100 (A) NAME/KEY: misc\_recomb  
101 (B) LOCATION:3..25  
102 (C) IDENTIFICATION METHOD: experimental  
103 (D) OTHER INFORMATION:/evidence= EXPERIMENTAL  
104 /standard\_name= "Multiple cloning site" /label= MCS  
105 /note= "A sequence specifying recognition sites  
106 for the restriction endonucleases: NlaIV, BstYI, BamHI, AlwI,  
107 MboI,  
108 DpnI, AflII, MseI, SspI, NsiI."  
110 (ix) FEATURE:  
111 (A) NAME/KEY: misc\_recomb  
112 (B) LOCATION:74..98  
113 (C) IDENTIFICATION METHOD: experimental  
114 (D) OTHER INFORMATION:/evidence= EXPERIMENTAL  
115 /standard\_name= "Multiple cloning site"  
116 /label= MCS  
117 /note= "A sequence specifying recognition sites  
118 for the restriction endonucleases: ScaI, RsaI, HpaI, HincII,

RAW SEQUENCE LISTING DATE: 02/12/2001  
PATENT APPLICATION: US/09/242,657B TIME: 10:52:56

Input Set : A:\Pto.amc  
Output Set: N:\CRF3\02122001\I242657B.raw

119 MseI, SfcI,  
120 PstI, Fnu4HI, BbvI, PvuII, NspBII, AluI, EcoRI."  
122 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
124 CGGGATCCTT AAGAATATTA TGCATNNNN AGTTTATTCT TGACANNNN NNNNNNNNNT 60  
126 GGTATAATAN NANAGTACTG TTAACTGCAG CTGAATTCCGG 100  
128 (2) INFORMATION FOR SEQ ID NO: 2:  
130 (i) SEQUENCE CHARACTERISTICS:  
131 (A) LENGTH: 113 base pairs  
132 (B) TYPE: nucleic acid  
133 (C) STRANDEDNESS: double  
134 (D) TOPOLOGY: linear  
136 (ii) MOLECULE TYPE: DNA (genomic)  
138 (iii) HYPOTHETICAL: YES  
140 (iv) ANTI-SENSE: NO  
142 (ix) FEATURE:  
143 (A) NAME/KEY: promoter  
144 (B) LOCATION: 23..95  
145 (D) OTHER INFORMATION:/standard\_name=  
146 "Artificial promoter library"  
147 /note= "A degenerated sequence specifying a mixture  
148 of artificial temperature regulated promoters covering a wide  
149 range of expression in small steps in L. lactis"  
151 (ix) FEATURE:  
152 (A) NAME/KEY: misc\_feature  
153 (B) LOCATION: 23..49  
154 (D) OTHER INFORMATION:/standard\_name=  
155 "Sequence providing temperature regulation to promoters"  
156 /note= "This sequence comprising two inverted  
157 repeats separated by a short spacer provides temperature (heat  
158 shock) regulation to promoters in Gram-positive bacteria"  
160 (ix) FEATURE:  
161 (A) NAME/KEY: misc\_feature  
162 (B) LOCATION: 50..60  
163 (D) OTHER INFORMATION:/standard\_name=  
164 "Consensus sequence"  
166 (ix) FEATURE:  
167 (A) NAME/KEY: misc\_feature  
168 (B) LOCATION: 75..84  
169 (D) OTHER INFORMATION:/standard\_name= "Consensus  
170 sequence"  
172 (ix) FEATURE:  
173 (A) NAME/KEY: misc\_feature  
174 (B) LOCATION: 89..95  
175 (D) OTHER INFORMATION:/standard\_name= "Consensus  
176 sequence"  
178 (ix) FEATURE:  
179 (A) NAME/KEY: -35\_signal  
180 (B) LOCATION: 55..60  
181 (D) OTHER INFORMATION:/standard\_name= "-35 box"

RECEIVED  
FEB 14 2001

TECH CENTER 1600/2900

DATE: 02/12/2001  
TIME: 10:52:56

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/242,657B  
Input Set : A:\Pto.amc  
Output Set: N:\CRF3\02122001\I242657B.raw

```

183     (ix) FEATURE:
184         (A) NAME/KEY: -10_signal
185         (B) LOCATION:78..83
186         (D) OTHER INFORMATION:/standard_name= "Pribnow box"
188     (ix) FEATURE:
189         (A) NAME/KEY: misc_recomb
190         (B) LOCATION:3..22
191         (D) OTHER INFORMATION:/standard_name= "Multiple
192 cloning site"
193 /label= MCS
194 /note= "A sequence specifying recognition sites
195 for the restriction endonucleases: NlaIV, BstYI, BamHI, AlwI,
196 MboI, DpnI, HindIII, AluI, MseI (2 sites), SspI, AseI."
198     (ix) FEATURE:
199         (A) NAME/KEY: misc_recomb
200         (B) LOCATION:89..111
201         (D) OTHER INFORMATION:/standard_name= "Multiple
202 cloning site"
203 /label= MCS
204 /note= "A sequence specifying recognition sites
205 for the restriction endonucleases: ScaI, RsaI, SfcI, PstI,
206 Fnu4HI, BbvI,
207 PvuII, NspBII, AluI, XbaI, MaeI, EcoRI, ApoI."
209     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
211 CGGGATCCAA GCTTAATATT AATTAGCACT CNNNNNNNNN GAGTGCTAAT TTTTTTGACA 60
213 NNNNNNNNNN NNNNTGGTAT AATANNANAG TACTGCAGCT GTCTAGAATT CGG      113
215 (2) INFORMATION FOR SEQ ID NO: 3:
217     (i) SEQUENCE CHARACTERISTICS:
218         (A) LENGTH: 199 base pairs
219         (B) TYPE: nucleic acid
220         (C) STRANDEDNESS: double
221         (D) TOPOLOGY: linear
223     (ii) MOLECULE TYPE: DNA (genomic)
225     (iii) HYPOTHETICAL: YES
227     (iv) ANTI-SENSE: NO
229     (vi) ORIGINAL SOURCE:
230         (A) ORGANISM: Saccharomyces cerevisiae
232     (ix) FEATURE:
233         (A) NAME/KEY: protein_bind
234         (B) LOCATION:10..16
235         (D) OTHER INFORMATION:/function= "Activating
236 promoters in
237 S. cerevisiae"
238 /bound_moiety= "GCN4 protein"
239 /standard_name= "Upstream activating sequence" /label=
240 UAS_GCN4p
241 /note= "A DNA sequence that specifies a binding
242 site for the GCN4 protein, which activates the transcription
243 of genes involved in amino acid synthesis in S. cerevisiae."

```

RAW SEQUENCE LISTING DATE: 02/12/2001  
PATENT APPLICATION: US/09/242,657B TIME: 10:52:56

Input Set : A:\Pto.amc  
Output Set: N:\CRF3\02122001\I242657B.raw

245 (ix) FEATURE:  
246 (A) NAME/KEY: TATA\_signal  
247 (B) LOCATION:67..72  
248 (D) OTHER INFORMATION:/standard\_name= "TATA box"  
250 (ix) FEATURE:  
251 (A) NAME/KEY: misc\_signal  
252 (B) LOCATION:122..144  
253 (D) OTHER INFORMATION:/function= "Transcription initiation"  
254 /standard\_name= "TI box"  
257 (ix) FEATURE:  
258 (A) NAME/KEY: protein\_bind  
259 (B) LOCATION:122..144  
260 (D) OTHER INFORMATION:/bound\_moiety= "Arginine repressor"  
261 /standard\_name= "arginine repressor binding site"  
264 /label= argR  
266 (ix) FEATURE:  
267 (A) NAME/KEY: misc\_RNA  
268 (B) LOCATION:145..192  
269 (D) OTHER INFORMATION:/function= "Spacer"  
270 /standard\_name= "Part of native sequence for ARG8 gene incl. first codon"  
274 (ix) FEATURE:  
275 (A) NAME/KEY: misc\_recomb  
276 (B) LOCATION:3..8  
277 (D) OTHER INFORMATION:/standard\_name= "Recognition site for restriction endonuclease EcoRI"  
279 /label= EcoRI\_site  
281 (ix) FEATURE:  
282 (A) NAME/KEY: misc\_recomb  
283 (B) LOCATION:192..197  
284 (D) OTHER INFORMATION:/standard\_name= "Recognition site or restriction endonuclease BamHI"  
286 /label= BamHI\_site  
288 (ix) FEATURE:  
289 (A) NAME/KEY: promoter  
290 (B) LOCATION:10..192  
291 (D) OTHER INFORMATION:/standard\_name= "Artificial promoter library"  
293 /note= "A degenerated sequence specifying a mixture of artificial promoters covering a wide range of expression in small steps in S. cerevisiae"  
298 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
300 CAGAATTCGT GACTCANNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN 60  
302 NNNNNNNNNN NNNNNNTATA AANNNNNNNN NNNNNNNNNN NNNNNNNNNN 120

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/242,657B

DATE: 02/12/2001

TIME: 10:52:57

Input Set : A:\Pto.amc

Output Set: N:\CRF3\02122001\I242657B.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:39 M:220 C: Keyword misspelled or invalid format, [(C) REFERENCE/DOCKET NUMBER:]